

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS

## **A decomposição em valores singulares**

TRABALHO DE CONCLUSÃO DE CURSO

**Paula Fernanda Vera**

FOZ DO IGUAÇU - SC

MARÇO - 2011

*Dedicatória:* Dedico a quem folhar, ler ou uma linha aproveitar.

*Agradecimentos:* Primeiramente ao Prof. Mario Zambaldi pelos ensinamentos, orientações e atenção dispendida que só assim tornaram possível a realização deste trabalho.

Aos meus colegas que durante o tempo de curso foram parceiros de estudos e aflições em véspera de provas.

Por fim, ao tutor Gilberto pelas orientações, preocupações e descontrações durante o curso.

*Mensagem:* "Se os teus projetos forem para um ano, semeia o grão. Se forem para dez anos, planta uma árvore. Se forem para cem anos, educa o povo" (Provérbio Chinês)

## *RESUMO*

Uma matriz arbitrária, mesmo não quadrada, pode ser fatorada por duas matrizes unitárias ortogonais e uma matriz contendo apenas uma submatriz diagonal, caracterizando assim uma generalização do teorema espectral. Esse tipo de fatoração é chamada de Decomposição em Valores Singulares, que dada a sua aplicabilidade é uma das fatorações mais importantes da Álgebra Linear. Neste trabalho mostraremos a sua existência e algumas de suas interessantes aplicações. e uma matriz contendo apenas uma submatriz diagonal, caracterizando assim uma generalização do teorema espectral.

# Sumário

Introdução	1
<b>1 Conceitos básicos da Álgebra Linear</b>	<b>3</b>
1.1 Os quatro espaços fundamentais de uma matriz . . . . .	3
1.2 Ortogonalidade . . . . .	7
1.3 Autovalores e autovetores . . . . .	9
1.4 Matrizes simétricas e o Teorema Espectral . . . . .	12
<b>2 A Decomposição em Valores Singulares</b>	<b>14</b>
2.1 Os valores singulares de uma matriz . . . . .	14
2.2 O Teorema da existência da SVD . . . . .	16
2.3 Algumas propriedades da $SVD$ . . . . .	21
2.4 Um procedimento para o cálculo da $SVD$ . . . . .	22
2.5 Bases ortogonais para os espaços fundamentais de uma matriz . . . . .	23
<b>3 Aplicações da SVD</b>	<b>25</b>
3.1 Sistemas lineares e o problema de quadrados mínimos . . . . .	25
3.2 A inversa generalizada de uma matriz . . . . .	28

3.3	Aplicação em processamento de imagens . . . . .	30
3.4	Exemplos da SVD com o MATLAB . . . . .	35
	<b>Conclusões</b>	<b>47</b>
	<b>Referências Bibliográficas</b>	<b>47</b>

# Introdução

Nas aplicações relevantes de várias áreas das ciências, os modelos matemáticos são representados por equações, cujos coeficientes dão origem a matrizes que devem ser manipuladas para obtenção da solução do modelo por meio de métodos numéricos. Vários métodos numéricos baseiam-se em fatorações de matrizes que tem por objetivo transformá-las em outras matrizes mais simples conservando algumas propriedades. Um exemplo bem conhecido é a Decomposição em Valores Singulares (Singular Value Decomposition - SVD). Em muitos cursos de graduação esse conceito não é abordado e quando o é, o fazem apenas superficialmente.

A decomposição em valores singulares é um dos resultados mais importantes da Álgebra Linear, tanto na área computacional quanto na teórica. Há inúmeras aplicações da SVD, dentre as quais se destacam o cálculo da pseudo inversa - na resolução de sistemas lineares e resolução do problema de quadrados mínimos -, a determinação do posto de uma matriz, e a compressão de imagens.

Este trabalho consiste num estudo da SVD e algumas de suas aplicações com a utilização do MATLAB [6]. A idéia é empregar uma linguagem simples e acessível com a introdução deste importante tema e suas relações com os conceitos básicos da Álgebra Linear: fatorações, subespaços de uma matriz, ortogonalidade e autovalores.

O Capítulo 1 está dividido em quatro seções que tem por objetivo exibir conceitos básicos da Álgebra Linear que servirão de alicerce para a teoria que será apresentada neste trabalho.

O segundo capítulo é talvez o mais importante para o *leitor de primeira viagem*

e isso se justifica já pelas duas primeiras seções. A primeira seção tem a função de definir o conceito de valores singulares para que na seção seguinte possa ser mostrada a existência da SVD para qualquer matriz. As seções seguintes deste mesmo capítulo tratam do cálculo da SVD e das bases ortogonais para os espaços fundamentais obtidas pela decomposição.

O Capítulo 3 é dedicado às aplicações citadas anteriormente, destacando principalmente a terceira seção, que talvez seja a seção que mais aproxima o leitor ao que há de interessante e vantajoso na aplicação desta teoria. O quarto e último capítulo traz basicamente alguns exemplos aplicados no MATLAB.

# Capítulo 1

## Conceitos básicos da Álgebra Linear

Neste capítulo vamos descrever alguns conceitos e resultados básicos da Álgebra Linear que estão muito relacionados com a SVD.

### 1.1 Os quatro espaços fundamentais de uma matriz

Entre os conceitos mais importantes da Álgebra Matricial, estão os espaços vectoriais fundamentais de uma matriz  $A$  e sua transposta  $A^T$  (conforme em [9]). Temos seis espaços, e são eles:

- o espaço-linha de  $A$
- o espaço-coluna de  $A$
- o espaço-nulo de  $A$
- o espaço-linha de  $A^T$
- o espaço-coluna de  $A^T$
- o espaço-nulo de  $A^T$

Lembre-se que os vetores-linha da matriz  $A$  são os vetores-coluna de  $A^T$ , assim como os vetores-coluna de  $A$  são os vetores-linha de  $A^T$ . Portanto, os espaços-linha e

coluna de  $A$ , são respectivamente, os espaços-coluna e linha de  $A^T$ . Isso faz com que os seis espaços citados acima se reduzam a apenas quatro espaços fundamentais associados a matriz  $A$ :

Dada uma matriz  $A_{m \times n}$  definimos:

- O espaço Coluna de  $A$ .

$$\mathcal{R}(A) = \{b \in \mathbb{R}^m \mid Ax = b; x \in \mathbb{R}^n\}$$

- O espaço Nulo ou Núcleo de  $A$ .

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

- O espaço linha de  $A$ .

$$\mathcal{R}(A^T) = \{y \in \mathbb{R}^n \mid A^T z = y; z \in \mathbb{R}^m\}$$

- O espaço Nulo ou Núcleo de  $A^T$ .

$$\mathcal{N}(A^T) = \{z \in \mathbb{R}^m \mid A^T z = 0; \}$$

Estes quatro subespaços de  $\mathbb{R}^m$  ou  $\mathbb{R}^n$ , estão associados à fatoração  $LU$  e às soluções do sistema retangular  $Ax = b$  ou, sistema homogêneo  $Ax - b = 0$  associado. De fato, podemos escrever o sistema  $Ax = b$  na forma de combinação linear:

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b,$$

e portanto, todos os  $b$ 's factíveis para o sistema linear são aqueles que estão em  $\mathcal{R}(A)$ . Então, para obter uma base de  $\mathcal{R}(A)$ , basta obter vetores independentes de  $\mathbb{R}^m$  para os quais o sistema seja compatível.

## Obtendo as bases pela fatoração LU

Como os outros espaços também apresentam relações deste tipo, vamos analisar a fatoração e solução de sistemas retangulares (veja em [4] [9]). Vamos fazê-lo através de um exemplo.

Considere a matriz  $A$  e sua fatoração abaixo:

$$A = \begin{bmatrix} 1 & -1 & -1 & 2 \\ 4 & -3 & -1 & 8 \\ 3 & -1 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 4 & 1 & & \\ 3 & 2 & 1 & \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 & 2 \\ & 1 & 3 & 0 \\ & & 0 & 0 \end{bmatrix} = LU$$

Para obter  $\mathcal{N}(A)$  buscamos as soluções de  $Ax = 0$ , que são as mesmas de  $Ux = 0$  pois, o determinante de  $L$  é não-nulo e por conseguinte  $L$  possui inversa. Resolvendo o sistema temos:

$$\mathcal{N}(A) = \{(-2, -3, 1, 0), (-2, 0, 0, 1)\}$$

Para obter  $\mathcal{R}(A)$  buscamos os possíveis  $b$ 's para que o sistema  $Ax = b$  tenha solução. Note que,

$$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix}$$

Logo, para um vetor  $b \in \mathbb{R}^3$  arbitrário temos:

$$L^{-1}b = L^{-1} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 4b_1 \\ b_3 + 5b_1 - 2b_2 \end{bmatrix}$$

Logo, a condição para a solvabilidade é:  $5b_1 - 2b_2 + b_3 = 0$  e  $\mathcal{R}(A)$  é um plano no  $\mathbb{R}^3$  ortogonal ao vetor  $(5, -2, 1)$  (veja a última linha de  $U$ ). Portanto,

$$\mathcal{R}(A) = \{(1, 4, 3), (-1, -3, -1)\}$$

Observe que escolhemos como vetores para compor a base de  $\mathcal{R}(A)$ , aqueles geradores das colunas correspondentes aos pivôs na eliminação. Isso foi proposital, para ressaltar o fato, já mencionado anteriormente, que a fatoração nos forneceria os subespaços. Na verdade, a explicação para tal fato está simplesmente baseada nos conceitos de combinação e dependência linear. São elas que também justificam o fato de que a base para o espaço-linha é formado por linhas não nulas da matriz  $U$ :

$$\mathcal{R}(A^t) = \{(1, -1, -1, 2), (0, 1, 3, 0)\}$$

Um caminho para obter  $\mathcal{N}(A^t)$ , pode ser através da fatoração de  $A^t$ . Ainda sobre a  $LU$  veja que para um vetor estar em  $\mathcal{N}(A^t)$ , este tem que satisfazer  $A^t z = 0$  ou ainda  $z^t A = 0$ . E estes são exatamente os vetores linha de  $L^{-1}$  correspondentes às linhas nulas de  $U$ , que neste exemplo é formado apenas por um vetor:

$$\mathcal{N}(A^t) = \{(5, -2, 1)\}$$

### O Teorema Fundamental

O que foi discutido anteriormente pode ser sintetizado no seguinte resultado:

Se  $A_{m \times n}$  tem posto  $r$  então:

- $\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(A^t)) = r$
- $\dim(\mathcal{N}(A)) = n - r$
- $\dim(\mathcal{N}(A^t)) = m - r$
- $\mathcal{R}(A^t)$  e  $\mathcal{N}(A)$  são complementos ortogonais em  $\mathbb{R}^n$ .
- $\mathcal{R}(A)$  e  $\mathcal{N}(A^t)$  são complementos ortogonais em  $\mathbb{R}^m$ .

## 1.2 Ortogonalidade

As transformações ortogonais apresentam uma vantagem importante sobre as transformações elementares. Na presença de erros de arredondamento há um controle natural sobre o crescimento das normas dos vetores obtidos pelas imagens das sucessivas transformações, o que não ocorre nas transformações de Gauss, por exemplo. Começaremos com algumas definições e propriedades de ortogonalidade.

**Definição 1** Os vetores  $a_1, a_2, \dots, a_n$  são ortogonais se

$$a_i^T a_j = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases}.$$

**Definição 2** Uma matriz quadrada  $A$  de ordem  $n$  é ortogonal se suas colunas formam um conjunto de vetores ortonormais.

*Caracterização de uma matriz ortogonal  $A$*

- $A$  tem colunas ortonormais
- $A$  tem linhas ortonormais
- $A^{-1} = A^T$
- $(Ax)^T(Ay) = x^T y$
- $\|Ax\| = \|x\|$

Essas propriedades podem ser obtidas diretamente observando que:

$$A^T A = I.$$

Vejamos o caso em que  $A$  é uma matriz arbitrária de ordem  $3 \times 3$ . Considerando  $a_j$  a coluna  $j$  de  $A$ , temos:

$$A^T A = \begin{pmatrix} -a_1^T - \\ -a_2^T - \\ -a_3^T - \end{pmatrix} \begin{pmatrix} | & | & | \\ a_1 & a_1 & a_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

Assim:

$$A^T A = A A^T = I \quad \text{ou} \quad A^{-1} = A^T.$$

Ainda,

$$(Ax)^T(Ay) = x^T A^T A y = x^T I y = x^T y$$

e

$$\|Ax\|^2 = (Ax)^T(Ax) = x^T x = \|x\|^2 \Rightarrow \|Ax\| = \|x\|.$$

**Exemplo 1.1** Seja,

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

logo,  $A^T A = A A^T = I$ . E, portanto  $A^T = A^{-1}$ .

**Exemplo 1.2** A matriz canônica para a rotação anti-horária do  $\mathbb{R}^2$  de um ângulo  $\theta$ , preserva norma e, de fato, é ortogonal:

$$A = \begin{pmatrix} \cos \theta & -\text{sen } \theta \\ \text{sen } \theta & \cos \theta \end{pmatrix} \quad \text{e} \quad A^T = A^{-1} = \begin{pmatrix} \cos \theta & -\text{sen } \theta \\ -\text{sen } \theta & \cos \theta \end{pmatrix}.$$

**Exemplo 1.3** O mesmo acontece com a permutação, que somente troca a ordem das componentes de um vetor:

$$Ax = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ x_1 \end{pmatrix}.$$

De fato,

$$A^T = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{e} \quad A^T A = A A^T = I.$$

Uma característica importante das matrizes ortogonais é que o sistema linear a elas associadas tem solução imediata. Em termos de combinação linear, os coeficientes são

conhecidos e, como veremos mais adiante, eles são a projeção unidirecional nas direções individuais dos vetores coluna da matriz:

$$Ax = b \Rightarrow A^T Ax = A^T b \Rightarrow x = A^T b$$

ou ainda,

$$b = \sum_{j=1}^n x_j a_j \Rightarrow x_j = b^T a_j.$$

As propriedades das transformações ortogonais são bastante exploradas em vários métodos na Álgebra Linear. Em muitos deles, a idéia é obter, a partir de um dado sistema de equações, outro sistema mais simples que preserve algumas características específicas. No caso do problema de autovalores, por exemplo, a idéia é obter sucessivas matrizes semelhantes a original, usando transformações ortogonais.

### 1.3 Autovalores e autovetores

O Problema de Autovalores e Autovetores de uma matriz, consistem em obter  $\lambda$  e  $x$  tal que:

$$Ax = \lambda x$$

onde os escalares  $\lambda$  são os *autovalores* e os vetores não nulos  $x$  são os seus respectivos *autovetores* associados da matriz  $A$ . O sistema de equações acima é não linear, mas bem específico. Então, estamos interessados em:

- Vetores  $x$  que não mudam de direção sob a ação da transformação linear associada a matriz  $A$ :
- Obter  $\lambda$  de modo que a matriz  $(A - \lambda I)$  seja singular, isto é:

$$\det(A - \lambda I) = 0$$

o Obter  $x$  de modo que  $x \in \mathcal{N}(A - \lambda I)$ :

$$(A - \lambda I)x = 0 \quad \text{ou} \quad Ax = \lambda x$$

A equação  $\det(A - \lambda I)$  recebe o nome de polinômio característico  $P(\lambda)$  da matriz. A raízes deste polinômio são os autovalores procurados que possibilitam o cálculo dos autovetores a eles associados.

### Exemplo 1.4

Seja,

$$A = \begin{pmatrix} 0 & 0 & -4 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{pmatrix}.$$

Então,  $P(\lambda) = \det(A - \lambda I) = (\lambda - 4)^2(2 - \lambda)$  é o polinômio característico de  $A$ . Os autovalores encontrados (raízes do polinômio) são:  $\lambda_1 = 4$  e  $\lambda_2 = 2$ . Agora iremos procurar os autovetores associados.

Se  $\lambda = 4$ , então

$$(A - \lambda I)x = \begin{pmatrix} -4 & 0 & -4 \\ 2 & 0 & 2 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Resolvendo este sistema obtemos  $(x_1, x_2, -x_1)$ . Dessa forma, os autovetores da matriz  $A$  associados a  $\lambda = 4$  são:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{e} \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Se  $\lambda = 2$ , então

$$(A - \lambda I)x = \begin{pmatrix} -2 & 0 & -4 \\ 2 & 2 & 2 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Resolvendo, obtemos  $(-2x_3, x_3, x_3)$  e o autovetor associado é:

$$v_1 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Se a matriz  $A$  é diagonalizável, ou seja, se apresenta um conjunto completo de  $n$  autovetores  $v$  independentes associados a  $n$  autovalores  $\lambda$  (iguais ou distintos), respectivamente. Então podemos escrever:

$$S = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \quad \text{e} \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

De forma que,

$$S^{-1}AS = \Lambda$$

Também temos:

$$AS = S\Lambda, \quad A = S\Lambda S^{-1}$$

A matriz citada no *Exemplo 1.3* é um exemplo de matriz diagonalizável.

É importante ressaltar que a diagonalização está relacionada aos autovetores (dependência linear), e a inversibilidade com os autovalores ( $\lambda_i \neq 0$ ).

Alguns resultados relevantes são oportunos lembrar:

- $Tr(A) = \sum a_{ii} = \sum \lambda_i$
- $det(A) = \prod \lambda_i$
- Autovetores correspondes a autovalores distintos são independentes
- Se  $Ax = \lambda x$  então  $\lambda^{-1}$  é autovalor de  $A^{-1}$  e  $\lambda^k$  é autovalor de  $A^k$
- Se  $A$  é diagonalizável então  $A^k$  também o é, ainda,

$$\Lambda^k = S^{-1}A^kS.$$

- Se  $A$  e  $B$  são diagonalizáveis, então compartilham a mesma matriz de autovetores  $S$ , se e somente se  $AB = BA$ .

## 1.4 Matrizes simétricas e o Teorema Espectral

Se a matriz  $A$  de ordem  $n$  é normal, ou seja,  $A = A^t$  ela é diagonalizável, seus autovetores são independentes, então podemos escrever:

$$Q^t A Q = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

em que  $\Lambda$  é a matriz onde a diagonal principal é formada pelos autovalores de  $A$  e  $Q$  a matriz obtida dos autovetores normalizados e em ordem correspondente à matriz  $\Lambda$ .

A fatoração acima ainda pode ser escrita como

$$A = Q \Lambda Q^t \quad \text{e, portanto} \quad A = \sum_{i=1}^n \lambda_i q_i q_i^t$$

Isto é o que diz o famoso teorema espectral para matrizes simétricas (veja em [4] [9]).

**Exemplo 1.5** Seja,

$$A = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 6 & 1 \\ 0 & 1 & 6 \end{pmatrix}.$$

Esta matriz é simétrica e possui os seguintes autovalores e autovetores:

$$\lambda_1 = -2, v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}; \quad \lambda_2 = 7, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}; \quad \lambda_3 = 5, v_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Observe que os autovetores são ortogonais, como confirma o Teorema espectral.

Podemos normalizá-los e assim obter a decomposição mencionada anteriormente:

$$q_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad q_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad q_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

Considerando os temas expostos neste primeiro capítulo, veremos que a *SVD* está intimamente relacionada às matrizes simétricas e portanto o Teorema Espectral tem um papel fundamental neste sentido. Veremos que embora a matriz original seja arbitrária, a *SVD* nos fornecerá entre outras coisas bases *ortogonais* para os quatro espaços fundamentais.

# Capítulo 2

## A Decomposição em Valores Singulares

Neste capítulo vamos descrever um dos mais importantes processos de fatoração de matrizes: a decomposição em valores singulares a qual chamaremos *SVD* (do inglês: Singular Value Decomposition).

Como veremos mais adiante, esta decomposição possui propriedades interessantes e aplicações importantes.

### 2.1 Os valores singulares de uma matriz

Dada uma matriz qualquer  $A$  de ordem  $m \times n$ , então as matrizes  $(A^T A)_{n \times n}$  e  $(A A^T)_{m \times m}$  são simétricas e, pelo Teorema Espectral, são diagonalizáveis por autovetores ortogonormais. Além disso, essas matrizes possuem os mesmos autovalores positivos.

Seja  $\lambda$  um dos autovalores de  $A^T A$  e  $v$  o seu autovetor correspondente com norma unitária. Então:

$$\begin{aligned}
0 \leq \|Av\|^2 &= (Av)(Av) = (Av)^T(Av) = v^T A^T Av \\
&= v^T(A^T Av) = v^T(\lambda v) = \lambda(v^T v) = \lambda(v \cdot v) \\
&= \lambda\|v\|^2 = \lambda.
\end{aligned} \tag{2.1}$$

Isso significa que os autovalores de  $A^T A$  além de serem todos reais também são todos não negativos e, portanto, podemos extrair a raiz quadrada positiva desses autovalores. De fato, as matrizes  $A^T A$  e  $AA^T$  são *simétricas semi-positivas definidas* quando  $\lambda_i \geq 0$  ou *simétrica positivas definidas* quando  $\lambda_i > 0$ .

**Definição 3** Se  $A$  é uma matriz  $m \times n$ , os valores singulares de  $A$  são as raízes quadradas dos autovalores positivos de  $A^T A$ , isto é,  $\sigma_i = \sqrt{\lambda_i}$ . É convenção ordenar os valores singulares de forma que  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ .

### Exemplo 2.1

Seja

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

então

$$AA^T = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

cujos autovalores não nulos são  $\lambda_1 = 3$  e  $\lambda_2 = 1$ . Portanto os valores singulares de  $A$  são:

$$\sigma_1 = \sqrt{3} \quad \text{e} \quad \sigma_2 = 1$$

Observe que os autovalores de  $A^T A$  são 3, 1 e 0 e conforme citado anteriormente:  $AA^T$  simétrica positiva definida e  $A^T A$  é simétrica semi-positiva definida.

Agora vamos à decomposição de uma matriz em valores singulares!

## 2.2 O Teorema da existência da SVD

**Teorema 1** *Toda matriz  $m \times n$  pode ser fatorada na forma*

$$A = UDV^T$$

onde  $U$  é ortogonal de ordem  $m$ ,  $V$  é ortogonal de ordem  $n$  e  $D_{m \times n}$  tem a forma:

$$D = \left( \begin{array}{cccc|c} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & 0 \\ & & & \sigma_r & \\ - & - & - & - & - \\ & & 0 & & 0 \end{array} \right)$$

em que  $r$  é o posto de  $A$  e  $\sigma_1 \geq \sigma_2 \geq \dots \sigma_{r-1} \geq \sigma_r > 0$ .

### Demonstração

Seja uma matriz  $A_{m \times n}$  com valores singulares  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$  e  $\sigma_{r+1}^2 = \sigma_{r+2}^2 = \dots = \sigma_n^2 = 0$ . Então, existem uma matriz ortogonal  $U_{m \times m}$ , uma matriz ortogonal  $V_{n \times n}$  e uma matriz  $D_{m \times n}$  de forma que

$$A = UDV^T, \text{ onde}$$

$$D = V_1^T A^T A V_1 = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \end{pmatrix}.$$

*Prova:* Considere a matriz simétrica  $A^T A$  com autovalores não-negativos

$$\sigma_1^2 \geq \sigma_2^2 \geq \dots \geq \sigma_r^2 > 0 \quad \text{e} \quad \sigma_{r+1}^2 = \sigma_{r+2}^2 = \dots = \sigma_n^2 = 0$$

associados à autovetores  $v_1, v_2, \dots, v_n$ . Como  $(A^T A)_{n \times n}$  é simétrica, então existe uma base ortonormal para  $R^n$  formada pelos autovetores da mesma. Logo, construiremos a matriz ortogonal  $V$  como sendo:

$$V = \left( V_1 \mid V_2 \right), \text{ onde } V_1 = \left( v_1 \ \dots \ v_r \right) \text{ e } V_2 = \left( v_{r+1} \ \dots \ v_n \right).$$

Seja,

$$\Delta_1 = V^T A^T A V = \left( V_1 \mid V_2 \right)^T A^T A \left( V_1 \mid V_2 \right) = \begin{pmatrix} V_1^T \\ \text{---} \\ V_2^T \end{pmatrix} A^T A \begin{pmatrix} V_1 \mid V_2 \end{pmatrix}.$$

Então,

$$\Delta_1 = \begin{pmatrix} V_1^T A^T A V_1 & \mid & V_1^T A^T A V_2 \\ \text{---} & \mid & \text{---} \\ V_2^T A^T A V_1 & \mid & V_2^T A^T A V_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & & & \mid & & \\ & \sigma_2^2 & & \mid & & \\ & & \ddots & \mid & 0 & \\ & & & \sigma_r^2 & \mid & \\ - & - & - & - & \mid & - \\ & & 0 & & \mid & 0 \end{pmatrix} = \begin{pmatrix} D_r^2 & \mid & 0 \\ - & \mid & - \\ 0 & \mid & 0 \end{pmatrix}$$

De fato, como  $\{v_1, \dots, v_n\}$  é uma base ortonormal de  $R^n$  então, sempre que  $i \neq j$ , temos

$$\begin{aligned} v_i^T A^T A v_j &= v_i^T (A^T A) v_j \\ &= v_i^T \lambda_j v_j \\ &= \lambda_j (v_i^T \cdot v_j) = 0. \end{aligned} \tag{2.2}$$

Logo, os blocos  $V_1^T A^T A V_2$  e  $V_2^T A^T A V_1$  são nulos. Iremos agora analisar os blocos restantes:

$$\begin{aligned}
 V_2^T A^T A V_2 &= v_i^T (A^T A) v_i, \quad \text{onde } r < i \leq n \\
 &= v_i^T \lambda_i v_i = \lambda_i (v_i^T \cdot v_i) \\
 &= \sigma_i^2 (v_i^T \cdot v_i) = 0 \cdot (v_i \cdot v_i) \\
 &= 0
 \end{aligned} \tag{2.3}$$

e,

$$\begin{aligned}
 V_1^T A^T A V_1 &= v_j^T (A^T A) v_j, \quad \text{onde } 1 \leq j \leq r \\
 &= v_j^T \lambda_j v_j = \lambda_j (v_j^T \cdot v_j) \\
 &= \sigma_j^2 (v_j \cdot v_j) = \sigma_j^2 \|v_j\|^2 \\
 &= \sigma_j^2.
 \end{aligned} \tag{2.4}$$

Portanto,

$$V_1^T A^T A V_1 = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r^2 \end{pmatrix} = D_r^2.$$

Pela equação (2.2) podemos observar que do fato de  $\{v_1, \dots, v_n\}$  ser uma base ortonormal decorre a ortogonalidade do conjunto de vetores  $\{Av_1, \dots, Av_n\}$  para  $R^m$ . Lembre-se que  $\sigma_i = \|Av_i\|$  e que  $\sigma_i$ , para  $i = 1, \dots, r$ , é não nulo. Portanto, podemos proceder da seguinte forma:

Se  $u_i = \frac{Av_i}{\sigma_i}$  então o conjunto  $\{u_1, \dots, u_r\}$  será ortonormal para  $R^m$ . No caso em que  $r < m$ , estendemos o conjunto á uma base ortonormal  $\{u_1, \dots, u_r, \dots, u_m\}$  de  $R^m$ .

Agora considere a matriz

$$U = \left( U_1 \mid U_2 \right),$$

em que  $U_1 = AV_1D_r^{-1}$  e  $U_2$  é a matriz formada pelos vetores do complemento ortogonal de  $U_1$ . Então,

$$U^T AV = \begin{pmatrix} U_1^T \\ \text{---} \\ U_2^T \end{pmatrix}^T A \begin{pmatrix} V_1 \mid V_2 \end{pmatrix} = \begin{pmatrix} U_1^T AV_1 & \mid & U_1^T AV_2 \\ \text{---} & \mid & \text{---} \\ U_2^T AV_1 & \mid & U_2^T AV_2 \end{pmatrix}$$

Vamos calcular as entradas de cada um dos blocos:

$$\begin{aligned} (11) \quad U_1^T AV_1 &= (AV_1D_r^{-1})^T AV_1 \\ &= (D_r^{-T}V_1^T A^T)AV_1 \\ &= D_r^{-T}(V_1^T A^T AV_1) \\ &= D_r^{-T}D_r^2 = D_r^{-1}D_rD_r = D_r. \end{aligned}$$

$$\begin{aligned} (12) \quad U_1^T AV_2 &= (AV_1D_r^{-1})^T AV_2 \\ &= (D_r^{-T}V_1^T A^T)AV_2 \\ &= D_r^{-T}(AV_1)^T AV_2 \\ &= D^{-T} \cdot 0 = 0. \end{aligned}$$

(21) Primeiramente, note que se  $U_1 = AV_1D_r^{-1}$  então  $U_1D_r = AV_1$ . Logo,

$$\begin{aligned} U_2^T AV_1 &= U_2^T(AV_1) \\ &= U_2^T(U_1D_r) \\ &= (U_2^T U_1)D_r \\ &= 0 \cdot D_r = 0 \end{aligned}$$

(22) Observe que  $AV_2 = 0$  pois, para todo  $i = r + 1, \dots, n$ ,  $\|Av_i\| = \sigma_i = 0$ . Seque então que

$$U_2^T AV_2 = U_2^T \cdot 0 = 0$$

Finalmente, note que

$$U^T AV = \left( \begin{array}{c|c} D_r & 0 \\ \hline - & - \\ 0 & 0 \end{array} \right) = D.$$

Portanto,

$$\begin{aligned} U^T AV &= D \\ (UU^T)A(VV^T) &= UDV^T \\ (UU^{-1})A(VV^{-1}) &= UDV^T \\ A &= UDV^T \end{aligned}$$

como queríamos mostrar.

**Exemplo 2.2:** Como no *Exemplo 2.1*,

se

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

então

$$U = \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad D = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{e} \quad V = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{-2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Em seguida vamos apresentar algumas propriedades interessantes da *SVD*

## 2.3 Algumas propriedades da $SVD$

A demonstração do Teorema da existência da  $SVD$  sugere alguns resultados importantes. Observando as matrizes  $A^T A$  e  $AA^T$ , o teorema espectral para matrizes simétricas fornece os resultados abaixo.

Como  $U$  é ortogonal,

$$A^T A = (UDV^T)^T(UDV^T) = VDU^TUDV^T = VD^2V^T$$

ou seja,  $\sigma_1^2, \dots, \sigma_r^2$  são os autovalores não nulos de  $A^T A$ .

Como  $V$  é ortogonal,

$$AA^T = (UDV^T)(UDV^T)^T = UDV^TVDU^T = UD^2U^T$$

ou seja,  $\sigma_1^2, \dots, \sigma_r^2$  são os autovalores não nulos de  $AA^T$ .

Um procedimento para obter a fatoração  $SVD$  pode ser desenvolvido baseado nessas relações, é o que apresentamos em seguida.

**Exemplo 2.3:** Novamente, recorrendo ao *Exemplo 2.1*

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

então

$$AA^T = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

cujos autovalores e autovetores (*já normalizados*) são:

$$\lambda_1 = 3, \quad u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \quad \lambda_2 = 1, \quad u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Ainda,

$$A^T A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

cujos autovalores e autovetores são:

$$\lambda_1 = 1, \quad v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}; \quad \lambda_2 = 0, \quad v_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ e } \quad \lambda_3 = 3, \quad v_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

Podemos observar que para calcular a  $SVD$  podemos calcular os autosistema de  $A^T A$  e  $AA^T$ . Dependendo das dimensões  $n$  e  $m$ , podemos calcular a *menor* delas ( $AA^T$  como no Exemplo 2.1  $m = 2$ ). Neste caso teremos os valores singulares  $\sigma'_i$ s e os vetores  $u'_i$ s. Para calcular os vetores  $v'_i$ s que faltam para compor a  $SVD$ , não precisamos calcular necessariamente  $A^T A$ . É isso que estabeleceremos a seguir.

## 2.4 Um procedimento para o cálculo da $SVD$

Baseado nas propriedades anteriores podemos calcular a  $SVD$  como se segue.

Como  $A = UDV^T$  então,

$$\begin{aligned} AV &= UD \quad \Rightarrow \quad Av_j = u_j \sigma_j \\ Av_j = u_j \sigma_j &\quad \Rightarrow \quad u_j = \frac{1}{\sigma_j} Av_j \quad \text{quando } j = 1, \dots, r \\ Av_j &= 0 \quad \text{quando } j = r + 1, \dots, n. \end{aligned}$$

Agora como  $A^T = VD^T U^T$  temos,

$$\begin{aligned} A^T U &= VD \quad \Rightarrow \quad A^T u_j = v_j \sigma_j \\ A^T u_j = v_j \sigma_j &\quad \Rightarrow \quad v_j = \frac{1}{\sigma_j} A^T u_j \quad \text{quando } j = 1, \dots, r \\ A^T u_j &= 0 \quad \text{quando } j = r + 1, \dots, m. \end{aligned}$$

Procedimento:

Caso a matriz  $A$  tenha mais linhas que colunas,  $m > n$ , faremos  $A^T A$  mas, se tiver mais colunas que linhas,  $m < n$ , faremos  $AA^T$ .

Calculamos os autovalores de  $A^T A$  ou de  $AA^T$  e depois determinamos os valores singulares de  $A$ .

**Exemplo 2.4:** Voltando ainda os caso do *Exemplo 2.1*, como  $m = 2$  e  $n = 3$  tendo calculado o autosistema de  $AA^T$ , faltaríamos calcular  $v_i$ 's por

$$v_i = \frac{1}{\sigma_i} A^T u_i$$

Assim

$$v_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$v_2 = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Finalmente  $v_3 \in N(A^T)$ , isto é,  $A^T u_3 = 0$  o que nos dá:

$$v_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

## 2.5 Bases ortogonais para os espaços fundamentais de uma matriz

A SVD nos fornece bases ortogonais para os quatro espaços fundamentais mencionados no Capítulo 1. Isso está baseado nas seções precedentes em que vimos que:

- Os vetores  $u_1 \dots u_r$  formam uma base ortogonal para  $R(A)$ :

$$Av_j = u_j \sigma_j \quad \text{quando } j = 1, \dots, r.$$

- Os vetores  $v_{r+1} \dots v_n$  formam uma base ortogonal para  $N(A)$ :

$$Av_j = 0 \quad \text{quando} \quad j = r + 1, \dots, n.$$

- Os vetores  $v_1 \dots v_r$  formam uma base ortogonal para  $R(A^T)$ :

$$A^T u_j = \sigma_j v_j \quad \text{quando} \quad j = 1, \dots, r.$$

- Os vetores  $u_{r+1} \dots u_m$  formam uma base ortogonal para  $N(A^T)$ :

$$A^T u_j = 0 \quad \text{quando} \quad j = r + 1, \dots, m.$$

O *Exemplo 2.1*, amplamente utilizado, ilustra as bases em  $\mathbb{R}^2$  e  $\mathbb{R}^3$ .

# Capítulo 3

## Aplicações da SVD

### 3.1 Sistemas lineares e o problema de quadrados mínimos

Considere  $A$  uma matriz de dimensão  $m \times n$  e o sistema linear  $Ax = y$ . Uma vez conhecida a SVD de  $A$ , torna-se simples encontrar a solução deste sistema. Sendo  $A = UDV^T$  temos:

$$Ax = b \quad \Rightarrow \quad UDV^T x = b \quad \Rightarrow \quad DV^T x = U^T b.$$

Seja  $y = V^T x$ , então

$$Dy = U^T b \quad \Rightarrow \quad x = Vy.$$

Não é incomum que o sistema linear  $Ax = b$  em questão seja inconsistente, neste caso nossa opção é procurar um  $x'$  que faça com que  $Ax'$  fique o mais próximo possível de  $b$ . Podemos pensar em  $Ax'$  como uma aproximação do  $Ax$  ideal, ou seja, de  $b$ . Referente ao produto interno euclidiano, desejamos que  $\|Ax' - b\|$  seja o menor possível, pois assim melhor será a aproximação para a solução.

O Problema de Quadrados Mínimos (PQM) é justamente encontrar um  $x'$  que torne  $\|Ax' - b\|$  o menor possível. O termo *quadrados mínimos* tem origem no fato de que se  $e = \|Ax' - b\|$ , que podemos chamar de *vetor-erro*, resulta em uma aproximação da solução. Então  $e = (e_1, \dots, e_n)$  pode ser minimizado por uma solução de mínimos

quadrados  $\|e\| = (e_1^2, \dots, e_m^2)^{\frac{1}{2}}$ , o que também minimiza  $\|e\|^2 = (e_1^2, \dots, e_m^2)$ .

O aspecto mais importante do PQM é que, independente do vetor  $x$  selecionado, o vetor  $Ax$  pertence necessariamente ao espaço coluna da matriz  $A$ , ou seja,  $Ax \in R(A)$ . Pois para resolver o problema de mínimos quadrados tendo em vista o espaço coluna de  $A$ , temos que para cada vetor  $x \in R^n$ ,  $Ax$  é uma combinação linear das colunas de  $A$ , ou seja,  $Ax$  varia sobre o espaço coluna de  $A$ .

Utilizando  $A^T$  quando  $A$  tem posto completo, obtemos uma forma de resolver o problema de quadrados mínimos através das equações normais:

$$A^T Ax = A^T b$$

obtida da relação de ortogonalidade entre o resíduo e as colunas de  $A$ :

$$A^T(b - Ax) = 0 \Rightarrow A^T b - A^T Ax = 0 \Rightarrow A^T Ax = A^T b$$

$$\text{ou } x = (A^T A)^{-1} A^T b.$$

**Exemplo 3.1** Considere o sistema  $Ax = b$ , em que  $b \notin R(A)$ . Vamos encontrar a solução do PQM quando:

$$A = \begin{pmatrix} 2 & -2 \\ 6 & 4 \\ -4 & 8 \end{pmatrix} \quad b = \begin{pmatrix} 8 \\ 2 \\ 6 \end{pmatrix}$$

*Resolução:*

Para evitar a inversão de  $A^T A$  no cálculo de  $x = (A^T A)^{-1} A^T b$ , iremos resolver o sistema de equações normais  $A^T Ax = A^T b$ .

$$A^T A = \begin{pmatrix} 2 & 6 & -4 \\ -2 & 4 & 8 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ 6 & 4 \\ -4 & 8 \end{pmatrix} = \begin{pmatrix} 56 & -12 \\ -12 & 84 \end{pmatrix}$$

$$A^T b = \begin{pmatrix} 2 & 6 & -4 \\ -2 & 4 & 8 \end{pmatrix} \begin{pmatrix} 8 \\ 2 \\ 6 \end{pmatrix} = \begin{pmatrix} 4 \\ 40 \end{pmatrix}$$

Logo, o sistema normal  $A^T Ax = A^T b$  para este caso é:

$$\begin{pmatrix} 56 & -12 \\ -12 & 84 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 40 \end{pmatrix}$$

Resolvendo o sistema encontramos a seguinte solução de mínimos quadrados:

$$x_1 = \frac{17}{195} \quad \text{e} \quad x_2 = \frac{143}{285}$$

A resolução do PQM pelas equações normais nem sempre é aconselhável. Vamos ilustrar esse fato com o próximo exemplo.

**Exemplo 3.2** Considere a matriz

$$A = \begin{pmatrix} 1 & 1 \\ 1 + \epsilon & 1 \\ 1 & 1 + \epsilon \end{pmatrix}.$$

Então,

$$A^T A = \begin{pmatrix} 3 + 2\epsilon + \epsilon^2 & 3 + 2\epsilon \\ 3 + 2\epsilon & 3 + 2\epsilon + \epsilon^2 \end{pmatrix}.$$

Suponha que  $\epsilon$  seja um número tão pequeno que  $1 + \epsilon$  possa ser representado exatamente em nosso computador, mas tal que  $1 + \epsilon + \epsilon^2$  seja arredondado para  $1 + \epsilon$ .

Então, fazendo o devido arredondamento a matriz  $A^T A$  calculada é

$$\begin{pmatrix} 3 + 2\epsilon & 3 + 2\epsilon \\ 3 + 2\epsilon & 3 + 2\epsilon \end{pmatrix}$$

que é singular e, portanto, o PQM não pode ser resolvido pelas equações normais, pois não encontraríamos a inversa de  $A^T A$ .

Em seguida mostraremos como usar a SVD para resolver o PQM em casos como este.

## 3.2 A inversa generalizada de uma matriz

Vimos no problema de mínimos quadrados que sob a hipótese da matriz  $A$  possuir posto completo, podemos recorrer a equação normal  $A^T Ax = A^T b$ . Na ausência dessa hipótese, utilizamos a inversa generalizada de Moore-Penrose ou, simplesmente, pseudo inversa. A pseudo inversa de uma matriz  $A$  arbitrária pode ser calculada a partir da decomposição de valores singulares. Se  $A = UDV^T$  então

$$A^+ = VD^+U^T$$

A matriz  $A^+$  é chamada de Moore-Penrose e satisfaz as seguintes propriedades:

- $AA^+A = A$
- $A^+AA^+ = A^+$
- $AA^+$  e  $A^+A$  são simétricas.

**Exemplo 3.3** Vamos calcular a pseudo-inversa do *Exemplo 2.1*:

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Então calculando a *SVD* obtemos:

$$U = \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad D = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{e} \quad V = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{-2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Então,

$$A^+ = VD^+U^T = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{-2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{-2}{3} & \frac{-1}{3} \\ \frac{1}{3} & \frac{-1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

Depois de obtida a  $SVD$  e a pseudo inversa, podemos então calcular a solução do  $PQM$  por meio do teorema a seguir.

**Teorema 2** A solução  $x^+$  do  $PQM$ ,  $\min_x \|Ax - b\|$  é dada por:

$$x^+ = A^+b = (VD^+U^T)b$$

onde,

$$D^+ = \left( \begin{array}{cccc|c} \sigma_1^{-1} & & & & \\ & \sigma_2^{-1} & & & \\ & & \ddots & & 0 \\ & & & \sigma_r^{-1} & \\ - & - & - & - & - \\ & & 0 & & 0 \end{array} \right).$$

*Prova:*

Seja  $A = UDV^T$  e  $Ax - b = 0$ . Então,

$$\|Ax - b\|^2 = \|UDV^T x - b\|^2 = \|U^T U D V^T x - U^T b\|^2 = \|D V^T x - U^T b\|^2.$$

Considerando  $y = V^T x$  e  $c = U^T b$ , temos

$$\|Ax - b\|^2 = \|Dy - c\|^2$$

$$\|Ax - b\|^2 = (\sigma_1 y_1 - c_1)^2 + (\sigma_2 y_2 - c_2)^2 + \dots + (\sigma_r y_r - c_r)^2 + c_{r+1}^2 + \dots + c_m^2.$$

Para se ter uma solução de norma mínima devemos ter

$$(\sigma_1 y_1 - c_1)^2 + \dots + (\sigma_r y_r - c_r)^2 = \sum_{i=1}^r (\sigma_i y_i - c_i)^2 = 0$$

isto é,  $y_1^+ = \frac{c_1}{\sigma_1}, \dots, y_r^+ = \frac{c_r}{\sigma_r}$ , ou na forma matricial:

$$y = D^+ c = D^+ U^T b,$$

ou ainda,

$$x^+ = Vy = VD^+U^T b = A^+b.$$



**Exemplo 3.4** Vamos considerar um caso em que  $b \notin R(A)$  e que  $A$  não tem posto completo. Vamos explorar este caso considerando um problema de *ajuste linear tridimensional* (veja [9])

Suponha que temos um conjunto de pontos  $(0, 0, 0)$ ,  $(0, 0, 2)$  e  $(1, 1, 2)$  e estamos interessados em obter um plano  $\pi : c + dt + ez$  que melhor se ajuste aos três pontos dados.

Note que é impossível que o plano passe pelos três pontos dados, pois se fosse o caso o sistema linear:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c \\ d \\ e \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$$

teria solução, o que não é o caso. Buscamos a solução  $x = A^+b = (c, d, e)$ .

Resolvendo o PQM encontramos a solução:  $c = 1$ ,  $d = 1/2$  e  $e = 1/2$ .

### 3.3 Aplicação em processamento de imagens

Atualmente a transmissão e o armazenamento de dados são de grande importância, bem como a garantia na confiabilidade da transmissão ou armazenamento destes dados, e a extensão dos mesmos. Pois, tanto numa situação quanto na outra, estamos interessados na redução de tempo e dinheiro. O detectar e corrigir erros fica por conta da teoria de códigos porém, para tratar da extensão dos dados devemos recorrer á outros conceitos. Se o dado em questão for uma imagem digital, um dos algoritmos mais eficientes é a decomposição em valores singulares, sendo essa uma das suas mais impressionantes aplicações.

A compressão de imagens utilizando a SVD, é uma compressão com perda de dados (lossy), entenda que os dados desprezados são redundantes, ou seja, perde-se a qualidade da imagem mas não se perde as informações essenciais. Por exemplo, imagine

que com uma máquina digital - que normalmente, gravam muito mais informações do que o olho humano detecta - você tire um retrato de alguém ao ar livre. O plano de fundo não é tão importante quanto a pessoa e seu rosto que possui muitos detalhes. A compressão da imagem trata de economizar alguns pixels de partes menos importantes (como o plano de fundo) e manter os detalhes essenciais na pessoa.

Agora vamos supor que a imagem descrita acima foi convertida em uma matriz  $A$  de posto  $r$ , podemos calcular a sua SVD,  $A = UDV^t$  e a decomposição na forma de produto externo

$$A = \sigma_1 u_1 v_1^t + \dots + \sigma_r u_r v_r^t.$$

Neste caso, cada parcela está sendo multiplicada por valores singulares  $\sigma_i$  (lembre-se que  $\sigma_1 \geq \dots \geq \sigma_r > 0$ ) portanto, quanto menor for o valor singular, menor será sua contribuição para a construção de  $A$ , ou seja, para a imagem. E assim identificamos os dados "irrelevantes" que poderão ser descartados. Ao invés de tomarmos todos os  $r$   $\sigma$ 's, iremos utilizar apenas os primeiros  $k$  (logo, os maiores), aproximando a matriz  $A$  por uma matriz

$$A_k = \sigma_1 u_1 v_1^t + \dots + \sigma_k u_k v_k^t$$

que irá gerar uma imagem com menor qualidade mas compreensível o quanto quisermos. Deste modo, serão transmitidos apenas os  $k$  primeiros valores singulares mais os  $k$  vetores  $u_i$  de dimensão  $m$  e os  $k$  vetores  $v_i$  de dimensão  $n$ , isto é,  $k + km + kn = k(1 + m + n) < m \cdot n$  números.

O algoritmo abaixo realizará no MATLAB o que acabamos de descrever para uma imagem em preto e branco. De forma resumida, o programa irá transformar a imagem em uma matriz  $A$  de ordem  $m \times n$ , calculará sua SVD e através da escolha de um posto  $k$  para a matriz, será obtida uma aproximação  $A_k$  para a imagem original.

```
imdir = 'diretório do arquivo';
```

```
imfile1 = 'nome do arquivo';
```

```
A = imread([imdir, imfile1]);
```

```
A=reshape(A,[m 3n]);      (em que  $m$  é o número de linhas e  $3n$  é o número de colunas
```

```

multiplicado por 3)
figure(1)
colormap(gray)
imagesc(A)
title('Imagem original')
axis image off
A=double(A);
(U,S,V)=svd(A);
k = um inteiro positivo não nulo
Ak=U1(:,1:k)*S1(1:k,1:k)*V1(:,1:k)';
figure(2);
colormap(gray);
imagesc(Ak);
title('Aproximação para k = o valor escolhido')
axis image off

```

O comando `imread` faz a leitura e conversão da imagem na matriz  $A$  de ordem  $m \times n \times 3$ . O comando `svd` no MATLAB não suporta matrizes 3D, por este motivo remodelamos a matriz com o comando `reshape` fazendo com que ela permaneça com o mesmo número  $m$  de linhas mas possua  $3n$  colunas. O comando `double` é necessário uma vez que as operações numéricas são realizadas apenas nessa classe. Consideremos agora a imagem digital abaixo de formato RGB:

A matriz dessa imagem é do tamanho  $480 \times 640$  pixels, portanto se quisermos trabalhar com esta imagem estaremos mexendo com  $480 \times 640 = 307200$  números, o que pode elevar o custo. O próximo algoritmo que veremos trata da compressão de uma imagem RGB (como a da *Figura 3.1*) sendo ele um pouco diferente do primeiro algoritmo. A matriz  $A$  de uma imagem colorida é multidimensional, representada por  $m$  linhas,  $n$  colunas e 3 páginas. Será necessário então fazer a paginação da matriz para assim calcular a SVD e fazer o truncamento de cada uma das 3 páginas.

Imagem original



Figura 3.1: Imagem Ferz2

```
imdir = 'diretório do arquivo';  
imfile1 = 'nome do arquivo';  
A = imread([imdir, imfile1]);  
figure(1)  
imagesc(A)  
title('Imagem original')  
axis image off  
A=double(A);  
A1=A(:,:,1);  
A2=A(:,:,2);  
A3=A(:,:,3);  
(U1,S1,V1)=svd(A1);  
(U2,S2,V2)=svd(A2);  
(U3,S3,V3)=svd(A3);  
k = um inteiro positivo não nulo  
Ak(:,:,1)=U1(:,1:k)*S1(1:k,1:k)*V1(:,1:k)';
```

```

Ak(:,:,2)=U2(:,1:k)*S2(1:k,1:k)*V2(:,1:k)';
Ak(:,:,3)=U3(:,1:k)*S3(1:k,1:k)*V3(:,1:k)';
figure(2)
imagesc(uint8(Ak))
title('Aproximação para k = o valor escolhido')
axis image off

```

Toda vez que quisermos encontrar uma aproximação distinta  $A_k$  para  $A$  devemos repetir a rotina:

```

k = um inteiro positivo não nulo
Ak(:,:,1)=U1(:,1:k)*S1(1:k,1:k)*V1(:,1:k)';
Ak(:,:,2)=U2(:,1:k)*S2(1:k,1:k)*V2(:,1:k)';
Ak(:,:,3)=U3(:,1:k)*S3(1:k,1:k)*V3(:,1:k)';
figure(2)
imagesc(uint8(Ak))
title('Aproximação para k = o valor escolhido')
axis image off

```

Observe o quadro de imagens abaixo que foi obtido com o segundo algoritmo utilizando diferentes valores para  $k$ .

Note que para  $k = 16$  a imagem já é bem compreensível, e para  $k = 32$  pode-se dizer que é bastante satisfatória, mas aos invés de 307200 números serão necessários apenas  $32(1 + 480 + 640) = 35872$ , isto representa uma redução de aproximadamente 88% do total de números a serem transmitidos. Podemos concluir que  $A_{32}$  é uma boa aproximação para a matriz  $A$  da imagem.

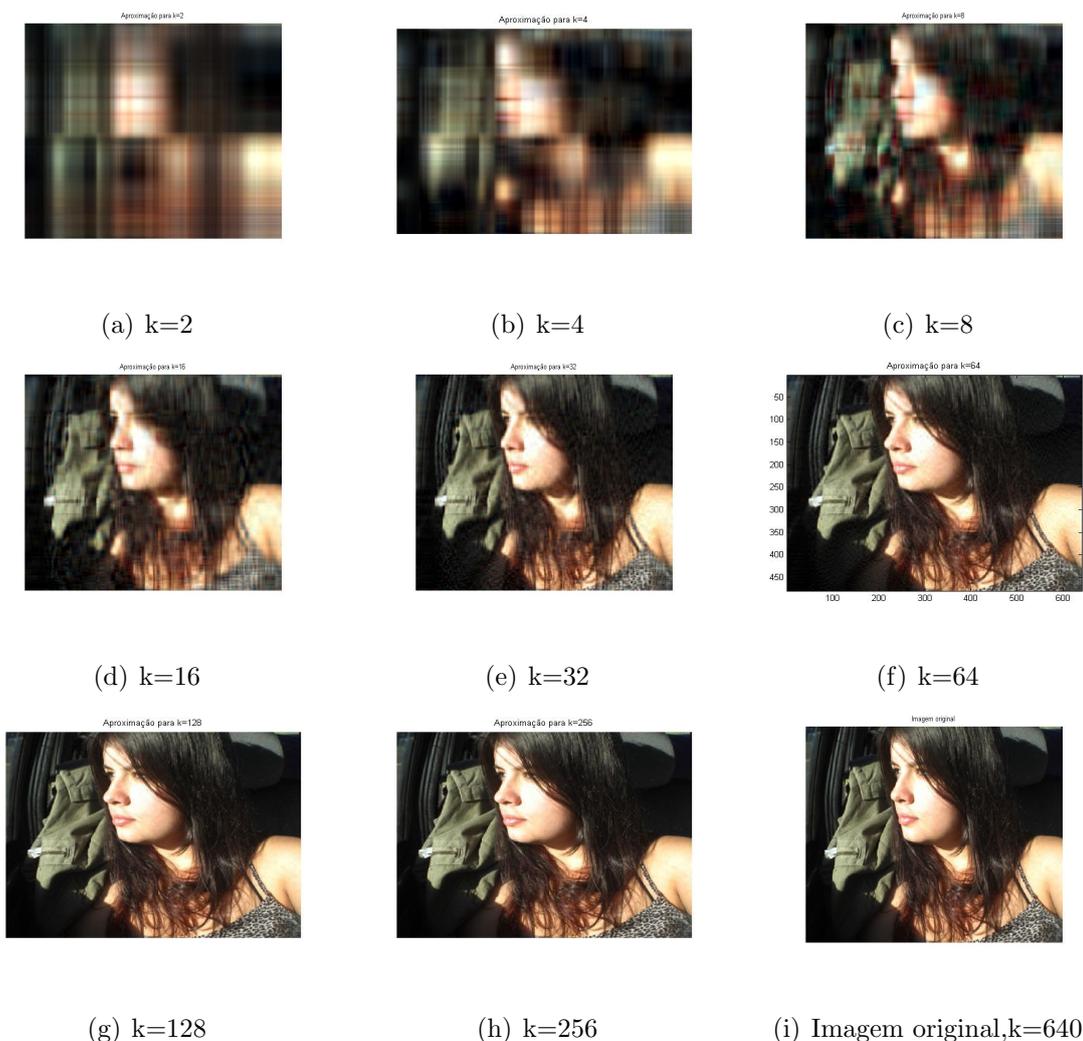


Figura 3.2: SVD da Imagem Ferz2

### 3.4 Exemplos da SVD com o MATLAB

Embora tenhamos calculado a  $SVD$  pelo autovalores das matrizes  $AA^T$  e  $A^T A$  como no capítulo anterior na prática isso não é aplicável. De fato, o mesmo ocorre com o cálculo dos autovalores, em que em problemas práticos, não podemos calcular autovalores tomando as raízes do polinômio característico por exemplo.

Nestes casos temos que lançar mão de métodos numéricos. De fato, existe um método específico no *MATLAB* para o cálculo da  $SVD$ . Uma vez no *MATLAB* lançamos o comando:

```
> [U,D,V]=svd(A)
```

e teremos a decomposição nas matrizes  $U, D, V$  em que  $D$  terá os valores singulares.

O procedimento usado no *MATLAB* é baseado em fatorações ortogonais. Trata-se do chamado método *QR* para obtenção de um tipo de *bidiagonalização* da matriz original. Este procedimento é descrito detalhadamente em [2] e está fora do escopo deste trabalho.

Agora vamos considerar alguns exemplos com o uso do *MATLAB*.

Nos três problemas que vamos considerar

$$m > n, \quad \text{posto}(A) < n \quad \text{e} \quad b \notin \mathcal{R}(A);$$

para melhor caracterizar um problema de quadrados mínimos.

Em todos os casos calculamos a *SVD* e a solução do Problema de quadrados mínimos  $x^+$  por meio da pseudo-inversa.

### Problema 1

$$m = 7, \quad n = 6, \quad \text{posto}(A) = 2$$

A =

3	6	9	12	15	18
6	9	12	15	18	21
9	12	15	18	21	24
12	15	18	21	24	27
15	18	21	24	27	30
18	21	24	27	30	33
21	24	27	30	33	36

b =

6.1119

13.2549

17.6922

21.9396

30.7239

35.0187

43.2591

D =

135.9958	0	0	0	0	0
0	9.4938	0	0	0	0
0	0	0.0000	0	0	0
0	0	0	0.0000	0	0
0	0	0	0	0.0000	0
0	0	0	0	0	0.0000
0	0	0	0	0	0

V =

-0.2567	0.6767	-0.3169	0.0051	-0.3182	-0.5239
-0.3128	0.4443	0.2759	-0.2611	-0.2460	0.7071
-0.3688	0.2119	0.5348	-0.0227	0.6743	-0.2789
-0.4249	-0.0204	-0.2723	0.8076	0.1309	0.2749
-0.4810	-0.2528	-0.5791	-0.5283	0.2902	0.0780
-0.5370	-0.4852	0.3576	-0.0006	-0.5312	-0.2572

U =

-0.2055	-0.6497	0.7100	0.1339	0.0800	-0.0827	0.0214
-0.2580	-0.4681	-0.4104	-0.3481	-0.1626	0.2267	0.5890
-0.3105	-0.2866	-0.3428	-0.1621	-0.2447	-0.5615	-0.5500
-0.3631	-0.1051	-0.1497	0.2500	0.0136	0.7367	-0.4789
-0.4156	0.0765	-0.1928	0.1036	0.8466	-0.2066	0.1189
-0.4681	0.2580	-0.0020	0.6282	-0.4221	-0.1880	0.3257
-0.5207	0.4396	0.3878	-0.6056	-0.1108	0.0755	-0.0261

$D^+ =$

0.0073	0	0	0	0	0	0
0	0.1053	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

$A^+ =$

-0.0459	-0.0329	-0.0198	-0.0068	0.0062	0.0193	0.0323
-0.0299	-0.0213	-0.0127	-0.0041	0.0045	0.0132	0.0218
-0.0139	-0.0098	-0.0056	-0.0014	0.0028	0.0070	0.0112
0.0020	0.0018	0.0016	0.0014	0.0011	0.0009	0.0007
0.0180	0.0134	0.0087	0.0041	-0.0006	-0.0052	-0.0099
0.0340	0.0249	0.0159	0.0068	-0.0023	-0.0113	-0.0204

$x^+ =$

1.0476

0.7619

0.4762

0.1905

-0.0952

-0.3810

## Problema 2

$$m = 10, \quad n = 4, \quad \text{posto}(A) = 2$$

$2A =$

2	4	6	8
4	6	8	10
6	8	10	12
8	10	12	14
10	12	14	16
12	14	16	18
14	16	18	20
16	18	20	22
18	20	22	24
20	22	24	26

$b =$

4.1794  
8.2464  
12.2846  
15.7610  
19.7936  
23.9372  
41.9700  
46.2720  
35.5040  
41.4660

D =

96.6008	0	0	0
0	5.3189	0	0
0	0	0.0000	0
0	0	0	0.0000
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

V =

0.4042	0.7325	-0.4059	-0.3678
0.4650	0.2895	0.2671	0.7929
0.5257	-0.1536	0.6835	-0.4825
0.5865	-0.5967	-0.5447	0.0573

U =

0.108	-0.577	-0.756	-0.000	-0.019	0.091	0.050	-0.082	0.252	-0.033
0.149	-0.475	0.417	0.169	0.373	0.196	0.240	0.122	0.118	0.532
0.190	-0.373	0.262	-0.195	-0.201	-0.320	-0.299	-0.663	-0.199	0.063
0.231	-0.271	0.132	0.145	-0.579	-0.310	-0.129	0.619	-0.027	-0.014
0.272	-0.168	0.066	-0.415	0.584	-0.191	-0.166	0.308	-0.132	-0.450

0.314	-0.066	0.151	-0.069	-0.239	0.828	-0.164	-0.019	-0.162	-0.266
0.355	0.035	0.034	-0.013	-0.133	-0.129	0.845	-0.147	-0.185	-0.257
0.396	0.137	-0.169	0.752	0.256	-0.093	-0.225	-0.096	-0.302	-0.049
0.437	0.239	0.159	0.035	-0.011	-0.074	-0.092	-0.134	0.823	-0.119
0.478	0.341	-0.293	-0.408	-0.031	0.002	-0.056	0.094	-0.183	0.595

$D^+ =$

0.0104	0	0	0
0	0.1880	0	0
0	0	0.0000	0
0	0	0	0.0000
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

$A^+ =$

-0.079	-0.064	-0.050	-0.036	-0.022	-0.007	0.006	0.020	0.034	0.049
-0.030	-0.025	-0.019	-0.013	-0.007	-0.002	0.003	0.009	0.015	0.020
0.017	0.014	0.011	0.009	0.006	0.003	0.000	-0.001	-0.004	-0.007
0.065	0.054	0.043	0.031	0.020	0.009	-0.001	-0.013	-0.024	-0.035

$\hat{x}^+ =$

1.8069

0.9958

0.1847

-0.6264

### Problema 3

$$m = 6, \quad n = 4, \quad \text{posto}(A) = 3$$

A =

12	28	4	8
12	28	4	8
0	16	4	0
0	16	4	0
8	4	0	0
8	4	0	0

b =

4
8
12
16
0
8

D =

50.4202	0	0	0
0	12.4072	0	0
0	0	4.8851	0
0	0	0	0.0000
0	0	0	0

0            0            0            0

V =

0.3388	0.8914	-0.2767	0.1185
0.9076	-0.3376	-0.0777	-0.2369
0.1472	-0.2322	-0.1618	0.9478
0.1993	0.1937	0.9440	0.1777

U =

0.6280	0.1502	0.2882	-0.6836	-0.1279	-0.1279
0.6280	0.1502	0.2882	0.6836	0.1279	0.1279
0.2997	-0.5103	-0.3871	-0.1809	0.4834	0.4834
0.2997	-0.5103	-0.3871	0.1809	-0.4834	-0.4834
0.1258	0.4659	-0.5168	0.0000	0.5000	-0.5000
0.1258	0.4659	-0.5168	0.0000	-0.5000	0.5000

D<sup>+</sup> =

0.0198	0	0	0
0	0.0806	0	0
0	0	0.2047	0
0	0	0	0.0000
0	0	0	0
0	0	0	0

A<sup>+</sup> =

-0.0013	-0.0013	-0.0127	-0.0127	0.0636	0.0636
0.0026	0.0026	0.0254	0.0254	-0.0022	-0.0022
-0.0105	-0.0105	0.0232	0.0232	0.0088	0.0088
0.0605	0.0605	-0.0816	-0.0816	-0.0921	-0.0921

$\hat{x}^+ =$

0.1368

0.7263

0.5947

-2.2947

# Conclusões

Em nosso trabalho, desenvolvemos um estudo sobre a Decomposição em Valores Singulares, suas propriedades, aplicações e experiências numéricas com o MATLAB. Ficamos surpresos em conhecer um conceito tão importante para a Matemática em si mesma como para as aplicações possíveis em outras áreas.

Ressalta-se também a importância da elaboração do capítulo destinado às aplicações como um meio de constatação da teoria vista.

Por tratar-se um tema que depende sobremaneira do conceito de autovetores e do Teorema Espectral, normalmente temas finais do curso de Álgebra Linear, raramente a decomposição SVD é vista no cursos de Licenciatura. Portanto, esse trabalho foi muito importante para revisão e estudos da parte básica e necessária dos assuntos relacionados. Esses estudos, assim como o da própria SVD, nos deram uma visão muito mais otimista da Álgebra Linear, uma disciplina muito mais importante do que considerávamos até então.

O texto foi construído com uma linguagem leve para que a leitura e a transmissão dos conceitos fosse agradável aos olhos dos alunos do curso que gostariam de aprofundar seus conhecimentos pós a graduação ou aos "curiosos" que apreciam as diversas vertentes da Álgebra Linear.

Por fim, o presente trabalho foi muito importante como complementação da minha formação, principalmente pelo contato direto com o ambiente de processamento numérico e computacional; e mais ainda para o reavivamento da vontade de continuar buscando novos saberes.

# Referências Bibliográficas

- [1] BJORCK, A. *Numerical Methods for Least Squares Problems*. Philadelphia: Siam, 1996.
- [2] GOLUB, G. H. AND CHARLES F. VAN LOAN. *Matrix Computations*.-2nd ed.- Baltimore: Johns Hopkins University Press,1989.
- [3] HIGHAM, N. J. *Accuracy and Stability of Numerical Algorithms*. Philadelphia: Siam, 1996.
- [4] LAY, D. C. *Álgebra linear e suas aplicações*. Rio de Janeiro: LTC, 1999.
- [5] LAWSON, C. L. AND RICHARD J. HANSON, R. *Solving Least Squares Problems*. Englewood Cliffs: N.J.: Prentice Hall, 1974.
- [6] LITTLEFIELD, B. *Matlab 6: curso completo*. São Paulo (SP): Prentice Hall, 2003.
- [7] NOBLE, B. AND JAMES W. DANIEL. *Applied Linear Algebra*.-3rd ed.- Englewood Cliffs, N. J.: Prentice Hall, 1988.
- [8] STEWART, G. W. *Introduction to Matrix Computations*. New Yor: Academic Press, Inc.,1973.
- [9] STRANG, G. *Linear Algebra and Its Applications*.-3rd ed.- San Diego: Harcourt Brace Jovanovich, 1988.
- [10] WATKINS, D. *Fundamentals of Matrix Computations*. New York: John Wiley & Sons, 1991.